Conformational Analysis of Three Pyrophosphate Model Species: Diphosphate, Methyl Diphosphate, and Triphosphate

MING-JING HWANG,[1] PEI-YING CHU,[1] JYE-CHAN CHEN,[2] ITO CHAO[2]

[1]*Institute of Biomedical Sciences*
[2]*Institute of Chemistry, Academia Sinica, 128 Yen-Chiou Yuan Rd., Sec. 2, Taipei, 11529 Taiwan, Republic of China*

Received 31 March 1999; accepted 21 July 1999

ABSTRACT: Whereas theoretical investigations of the energetic origin of hydrolyzing a pyrophosphate linkage abound, few studies have focused on the energetics of the rotation of this linkage. This less-studied property of the pyrophosphate linkage was investigated here by use of *ab initio* calculations to characterize the conformational space of three model species of pyrophosphate anions: diphosphate ($P_2O_7^{-4}$), methyl diphosphate ($CH_3P_2O_7^{-3}$), and triphosphate ($P_3O_{10}^{-5}$). By carefully selecting conformationally distinct rotational isomers of the three model compounds, their potential surfaces were thoroughly explored. In addition to showing that a terminal phosphate group is indeed very flexible in accordance with the general perception of free rotation, a number of intriguing features of this linkage emerged from the *ab initio* calculations, which include an influential sp3-hybridized C—H. . .O intramolecular hydrogen bond in methyl diphosphate, and a highly restricted rotational space pertaining to the central pyrophosphate linkage of the triphosphate anion. These *ab initio* findings were then evaluated by, and proved insightful in, follow-up examinations of experimentally determined complex structures of proteins and their dinucleotide or adenine and guanine triphosphate ligands.
J Comput Chem 20: 1702–1715, 1999

Keywords: pyrophosphate linkage; potential energy surface; *ab initio*; P—O—P angle; C—H. . .O hydrogen bond

Correspondence to: M.-J. Hwang; e-mail: mjhwang@ibms.sinica.edu.tw

Contract/grant sponsor: National Science Council; contract/grant number: NSC86-8211-B-001-0022

This article includes Supplementary Material available from the authors upon request or via the Internet at ftp.wiley.com/public/journals/jcc/suppmat/20/1702 or http://journals.wiley.com/jcc/

CCC 0192-8651 / 99 / 161702-14

Introduction

One of the most important chemical linkages in biology is the pyrophosphate (P—O—P) linkage of adenine and guanine di- and triphosphates (i.e., ADP, ATP, GDP, and GTP). It is through hydrolysis of this linkage and the reversed reaction that these nucleotides mediate numerous biological functions.[1] For this reason, for many years model species comprising the pyrophosphate linkage have been the subject of intensive theoretical studies.[2–9] However, the efforts have primarily been directed toward understanding the origin of the hydrolysis energies by which the term "energy rich" or "high energy"[10] has become a frequent annotation to this particular class of nucleotides. As to bond rotation, based primarily on the extended distribution of torsional states observed in crystal structures of various pyrophosphates,[11–13] a widely held perception is that this linkage is generally free to rotate.

In the present work, we evaluated this "free rotation" notion by exploring the potential energy surfaces of three pyrophosphate prototypes, diphosphate (DP, $P_2O_7^{-4}$), methyl diphosphate (MDP, $CH_3P_2O_7^{-3}$), and triphosphate (TP, $P_3O_{10}^{-5}$), using, primarily, HF/6-31G* *ab initio* calculations. To our knowledge, previous quantum mechanical calculations on these three molecules, shown in Figure 1, have been sketchy, based only on minimum-energy structures[9] or on semiempirical theories.[3, 14, 15] The most recent work on this subject is that of Pavelites et al.,[16] who, in deriving a set of empirical force-field parameters for the pyrophosphate linkage, reported a number of MDP torsional conformations optimized with HF/6-31+G*. As presented and discussed herein, our extensive calculations (over a hundred distinct rotational isomers of the three model compounds were calculated) showed that there is still much to know about the potential surface of the pyrophosphate rotation, which, as implied by the notion of free rotation, is thought to be very flat. In addition, where possible we evaluated implications of these *ab initio* results by comparing them to experimentally observed nucleotide conformations.

Methods

Through projections along adjacent phosphorus–phosphorus (P–P) axes as well as a phosphorus–carbon (P–C) axis for MDP, rotational isomers of the three model molecules can be classified by three

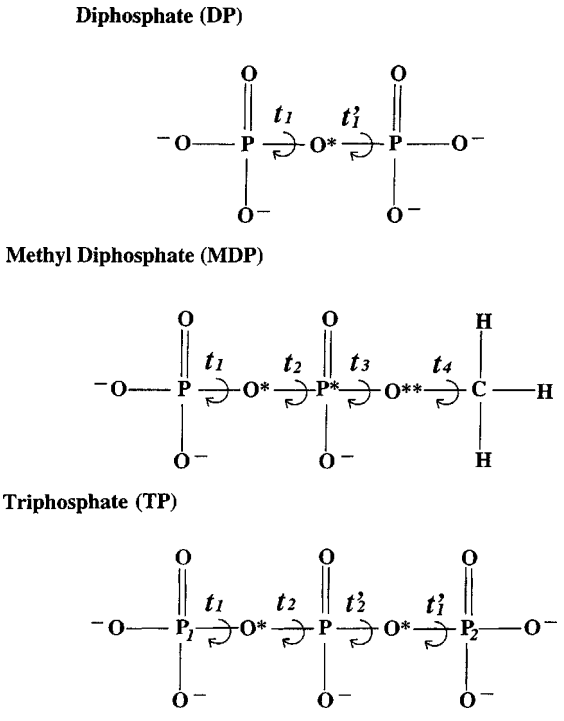


FIGURE 1. Schematic representations of the three model compounds. Torsions and key atoms of the pyrophosphate backbone are labeled.

groups: staggered (*stg*), skewed (*skw*), and eclipsed (*eclp*), and within each group the orientation of the bridge oxygen(s) (O* as well as O** in the case of MDP; Fig. 1) further distinguishes specific rotational conformations. These projections are depicted in Figure 2 for MDP, Figure S1 for DP, and Figure S2 for TP (throughout the article the figures and tables that are provided as supplemental materials are indicated by an "S" in front of their sequential number). Note that each of these projections goes through two atoms connected not by one bond but by two, and therefore, an *eclp* structure does not necessary possess a torsion angle of 0°, for example; rather, it is the combination of two torsions that determines the group type of these rotational isomers.

Using these projections, distinct rotational isomers for each of the three model molecules were generated by systematically incrementing, by 30° each time, torsions of the pyrophosphate linkage (Fig. 1) whose specific values were kept constant during the optimization. For a few minimum-energy structures, full-degree optimization followed by calculation of analytical second derivatives were also performed. All these calculations were carried out using Gaussian94[17] with HF/6-31G*, where polarization functions were added to phosphorus, oxygen, and carbon atoms.

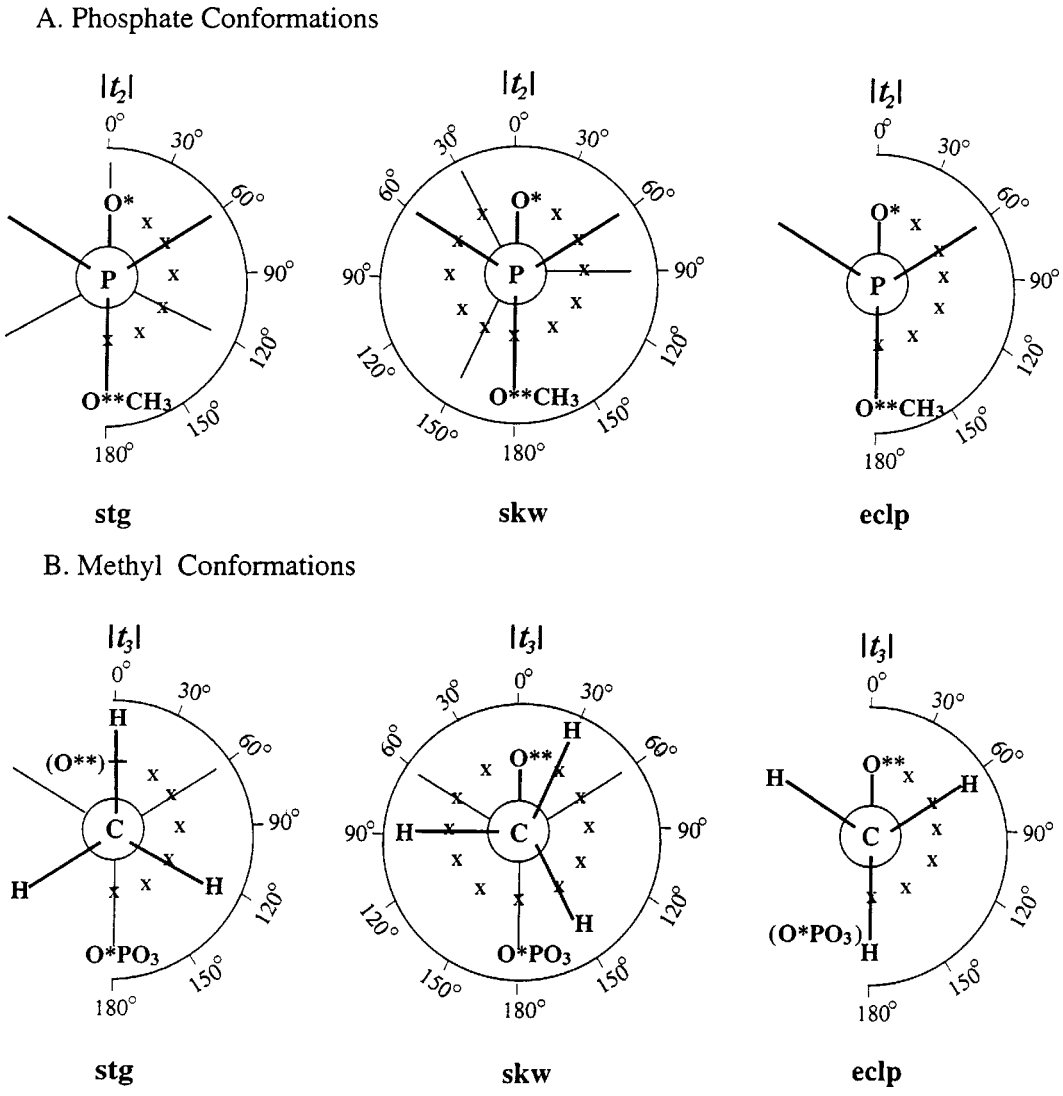


FIGURE 2. Projection of MDP rotational isomers for (A) the terminal phosphate group (P—P projection), (B) the methyl group (P—C projection). Each "x," which denotes the position of the bridge oxygen (O* or O**), represents a distinct rotational isomer.

The effects of a larger basis set and electron correlation were evaluated by three higher levels of *ab initio* calculations, MP2/6-31G*, HF/6-31+G*, and MP2/6-31+G*, on DP structures. The MP2 calculations were carried out without freezing core electrons, and diffuse functions were added to phosphorus, oxygen, and carbon atoms when the basis set 6-31+G* was used.

To evaluate the effect of solvation on conformational energy differences, four TP structures assuming the pyrophosphate torsional angles of ATP as respectively observed in four high-resolution crystal of protein–ATP complexes were calculated in both the gas phase and the water-solvated phase. For the solvation calculations, a single-point SCRF (self-consistent reaction field)[18] treatment was carried out with the program Jaguar[19] on the gas phase-optimized structures (by HF/6-31G*). In this continuum treatment, the dielectric constant of water was set at 80.37.

Results

DIPHOSPHATE (DP)

A total of 144 (t_1, t_1') rotational isomers of DP would be generated by a 30° torsional search for both t_1 and t_1', which denote, respectively, rotations of the two phosphate groups (Fig. 1). However,

TABLE I.

Conformational Energy Differences (ΔE in kcal/mol) and Selected Bond Lengths (Å) and Angles (deg) of the Seven Distinct Diphosphate (DP) Rotational Isomers[a] as Optimized by Four Different Levels of *ab initio* Calculations.

	Method[b]	P—O*—P	P—O*	ΔE[c]
stg				
I	A	**159.3**	**1.6668**	0.00
	B	152.0	1.7196	0.00
	C	168.3	1.6470	0.00
	D	156.6	1.6974	0.00
II	A	159.5	1.6667[d]	0.00036
	B	152.4	1.7159[d]	0.02165
	C	168.4	1.6468[d]	0.00110
skw				
I	A	160.2	1.6674[d]	**0.51**
	B	152.6	1.7209[d]	0.41
	C	170.5	1.6472[d]	0.49
	D	156.8	1.6989, 1.6986	0.37
II	A	162.9	1.6658[d]	**0.58**
	B	156.3	1.7184[d]	0.73
	C	**180.0**	1.6451	0.51
eclp				
I	A	162.3	1.6675	**1.09**
	B	153.6	1.7220	0.90
	C	**180.0**	1.6463	0.99
	D	157.5	1.6997	0.82
II	A	165.6	1.6642	**1.15**
	B	156.3	1.7205	1.09
	C[e]	**180.0**	—	—
III	A	168.3	1.6642	**1.18**
	B	158.9	1.7179	1.23
	C[e]	**180.0**	—	—

[a] See Figure 1 for the labeling of atoms, and Figure S1 for specific structures.
[b] A: HF/6-31G*, B: MP2/6-31G*, C: HF/6-31+G*, and D: MP2/6-31+G*.
[c] With respect to the minimum energy at each method of calculation.
[d] Two symmetry-nonequivalent P—O* bonds were computed to be of the same length.
[e] The same structure as C of *eclp* I.

due to high symmetry redundancy, only seven of them—two each for *stg* and *eclp*, and three for *skw*—are distinct, as Figure S1 illustrates. In Table I selected geometric parameters and relative energies of the seven distinct DP rotational isomers resulting from the four different levels of *ab initio* calculations are compared. Figure S3 depicts the whole (t_1, t_1') energy surface of DP, which was created by applying symmetry for equivalent conformations from these seven distinct DP structures.

The effects of a larger basis set (6-31+G* vs. 6-31G*) and inclusion of electron correlation (MP2 vs. HF) on the optimized geometries of DP can be seen in Table I. They appear to work in the opposite direction; namely, the pyrophosphate P—O* bond is shortened with the larger basis set but lengthened with electron correlation. Furthermore, to mitigate the increased repulsion between the two phosphate groups brought closer by a shorter P—O* bond, the P—O*—P angle is correspondingly more open in structures calculated with the method using the larger basis set, HF/6-31+G*, and without the correction of MP2. In fact, with HF/6-31+G* the P—O*—P angle is pushed to its limit (180°) in congested conformations such as *skw* II and all of the three *eclp* structures (see Table I). In combination, the two effects cancel each other out to a large extent, giving rise to, albeit fortuitously, a comparatively good agreement between HF/6-31G* and the highest level of the four methods employed (MP2/6-31+G*) in the calculated geometries. By contrast, no significant variations were observed in the conformational energy differences resulting from the four different levels of *ab initio* calculations. Previously, HF with 6-31G*[9, 20, 21] and basis sets of similar quality[4] have been shown to give satisfactory structures and energetics for phosphate compounds. All together, it seems that the HF/6-31G* level of calculations would be reasonable in describing conformational structures and energies of the pyrophosphate compounds to be studied here. Accordingly, unless noted otherwise, all the values discussed hereafter have been obtained with HF/6-31G*.

For DP, all of its seven distinct rotational isomers (Fig. S1) were also subjected to full-degree optimization by HF/6-31G*. The results show that without torsion constraints they all reverted to *stg* I, except *stg* II, which remained unchanged. Analysis of the second derivatives of the energy surface reveals that whereas *stg* I is an energy minimum of C_2 symmetry, *stg* II, with its bridge oxygen (O*) eclipsing one of the six terminal oxygens, is a maximum of C_2 symmetry. The minimum energy structure, *stg* I, obtained here is identical to that reported by Colvin et al.,[9] which is also a HF/6-31G* structure.

METHYL DIPHOSPHATE (MDP)

To locate energy minima of MDP, structures of differing conformations representing *trans* (*t*), *gauche* (*g*), and *cis* (*c*) for its central two torsions,

TABLE II.

Conformational Energy Differences (ΔE, in kcal/mol) and Selected Geometric Parameters[a] (Bond in Å, Angle and Torsion in Degree) for the Fully Optimized Minimum-Energy Structures of Methyl Diphosphate (MDP) Calculated by HF/61-31G*.

Conf.	(t_2, t_3)	P—O*—P*	P—O*	P*—O*	P*—O**	C—H...O[b]	ΔE
gg I	(84, −68)	149.3	1.7608	1.5635	1.6799	2.15	0.00
cg	(−27, −57)	150.6	1.7600	1.5634	1.6693	2.19	0.26
gg II	(50, −54)	149.4	1.7625	1.5649	1.6722	2.40	0.99
gt	(87, −168)	148.5	1.7894	1.5438	1.7023	2.60[c]	6.55

[a] See Figure 1 for the labeling of atoms and torsions.
[b] The shortest C—H...O distance.
[c] Where the oxygen belongs to the phosphate group adjacent to the methyl (see Fig. S4).

t_2 and t_3 (Fig. 1), respectively, were fully optimized. Four stable conformations were identified, which are shown in Figure S4. Their relative energies and key geometric parameters are given in Table II, whereas a complete list of the geometric parameters is provided in Table S1. They include two *gg* (I and II), one *gt*, and one *cg* structures. Attempts to obtain fully optimized *tg* and *tt* structures were not successful because a *trans* t_2 torsion linearized the P—O*—P* angle (also see below the results from torsion-constrained optimizations). The lowest-energy structure is *gg* I, which has a staggered phosphate conformation and torsion angles of 84° and −68°, respectively. Nearby there is a second *gg* minimum (*gg* II) of ∼1 kcal/mol higher in energy with slightly different torsion values, (50°, −54°), and an *skw* phosphate conformation. The *cg* structure, despite a *cis*-like torsion in t_2 (−27°), is more stable than *gg* II, and only 0.26 kcal/mol less stable than *gg* I. The *gt* structure, with ΔE of 6.55 kcal/mol, is a high-energy minimum. As discussed later, a combination of two intramolecular forces, the "*gauche* effect" and C—H...O hydrogen bonding, is primarily responsible for the differential stabilities of these MDP structures.

To further characterize the potential energy surface of MDP, additional structures representing, separately, phosphate and methyl conformations, were optimized at fixed torsion values. In the former, rotational isomers of the phosphate group, i.e., (t_1, t_2) torsions, were sampled, while in the latter the conformational space of (t_3, t_4) was examined (see Fig. 1). As may be seen from the projections shown in Figure 2(A) and 2(B), each of the two types of MDP conformations encompasses 26 distinct rotational isomers, with 7, 7, and 12 from the group *stg*, *eclp*, and *skw*, respectively. However, in the case of phosphate conformations, only 21 turn out to be distinct, because when t_2, defined by the dihedral angle of P—O*—P*—O** (where O** denotes the ester oxygen to which the methyl group is attached), is forced to take a *trans* (close to 180°) conformation, the P—O*—P* angle tends to be linearized. Additional runs starting with differing geometries were carried out to ensure that the linearization was not a result of an instability in structural optimization. The consequence of P—O*—P* linearization is an ill-defined t_2 torsion and degeneracy in some of these structures. In increasing order of relative energies within each conformational group, results of these torsion-constrained optimizations are presented in Table S2 for the phosphate conformations and in Table S3 for the methyl conformations. In Figure 3(A) and 3(B) the complete energy surface of these two types of MDP rotational conformations is constructed from the ΔE data of Tables S2 and S3, respectively.

TRIPHOSPHATE (TP)

To explore the conformational space of TP, it helps to think of TP as two DP fused on a common phosphate. Viewing TP as if viewing two DP separately, i.e., projecting the second phosphorus atom onto the first and the third separately (Fig. S2), one can divide the rotational isomers of TP into six groups: *stg–stg*, *stg–skw*, *stg–eclp*, *skw–skw*, *skw–eclp*, and *eclp–eclp*, with *skw–stg* structures being included in group *stg–skw*, *eclp–stg* structure in group *stg–eclp*, and so on. The six groups thus defined implicitly place restraints on the values of the terminal torsion, t_1 and t_1', within each group, and leave the central two torsion, t_2 and t_2', as the independent variables to search for distinct TP rotational isomers (see Fig. 1 for the labeling of these torsions). Furthermore, because the two terminal phosphates are interchangeable, rotational conforma-

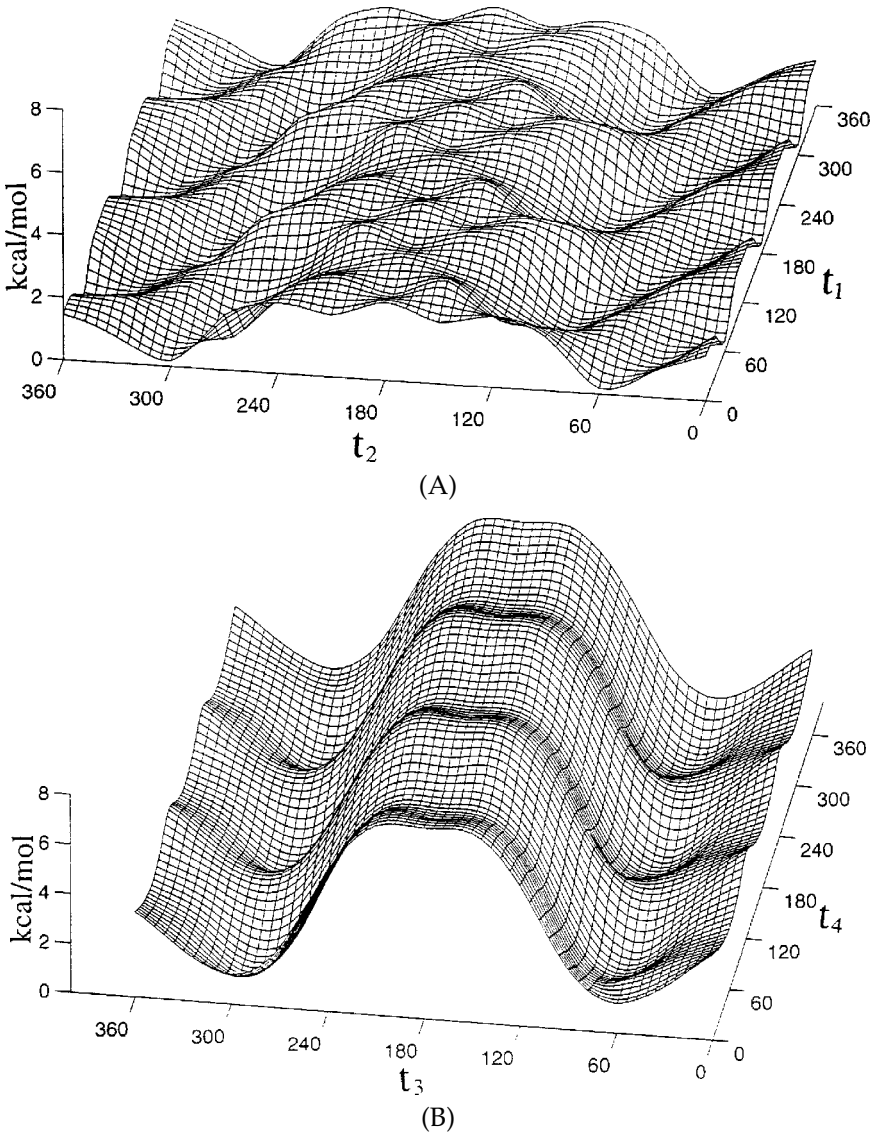


FIGURE 3. MDP potential energy surfaces (HF/6-31G* results): (A) phosphate conformations, (B) methyl conformations.

tions of (t_1, t_2, t'_2, t'_1), (t'_1, t'_2, t_2, t_1), $(-t_1, -t_2, -t'_2, -t'_1)$, and $(-t'_1, -t'_2, -t_2, -t_1)$ are equivalent. That is, within each conformational group, only a quadrant of the (t_2, t_2) surface needs to be considered. This is illustrated in Figure 4, where the elements of the (t_2, t'_2) conformational plane are shown to be symmetrical with respect to the two diagonal lines, $t_2 = t'_2$ and $t_2 = -t'_2$. Using an increment of 30°, the number of grid points for a quadrant of the (t_2, t'_2) plane is 49 (Fig. 4). Thus, for a one-to-one correspondence that holds for the two highly symmetric groups, *stg–stg* and *eclp–eclp*, there would be 48 distinct rotational isomers (note that (0°, 0°) is shown twice in the grid). However, for the other four less symmetric groups, because multiple distinct rotational isomers can possess identical (t_2, t'_2) torsions, their number will be larger than 48. By a careful examination, we obtained 123, 198, 84, and 123 as the number of distinct TP rotational isomers for group *stg–skw*, *skw–skw*, *stg–eclp*, and *skw–eclp*, respectively. Together with groups *stg–stg* and *eclp–eclp*, they add up to a total of 624 distinct TP rotational isomers.

On the other hand, due to a stronger electrostatic repulsion, folded TP structures are much susceptible to P—O—P linearization, even at the level of HF/6-31G* calculations. This is also shown in Figure 4, where the shaded area denotes conformations at which structure optimization would lead to a

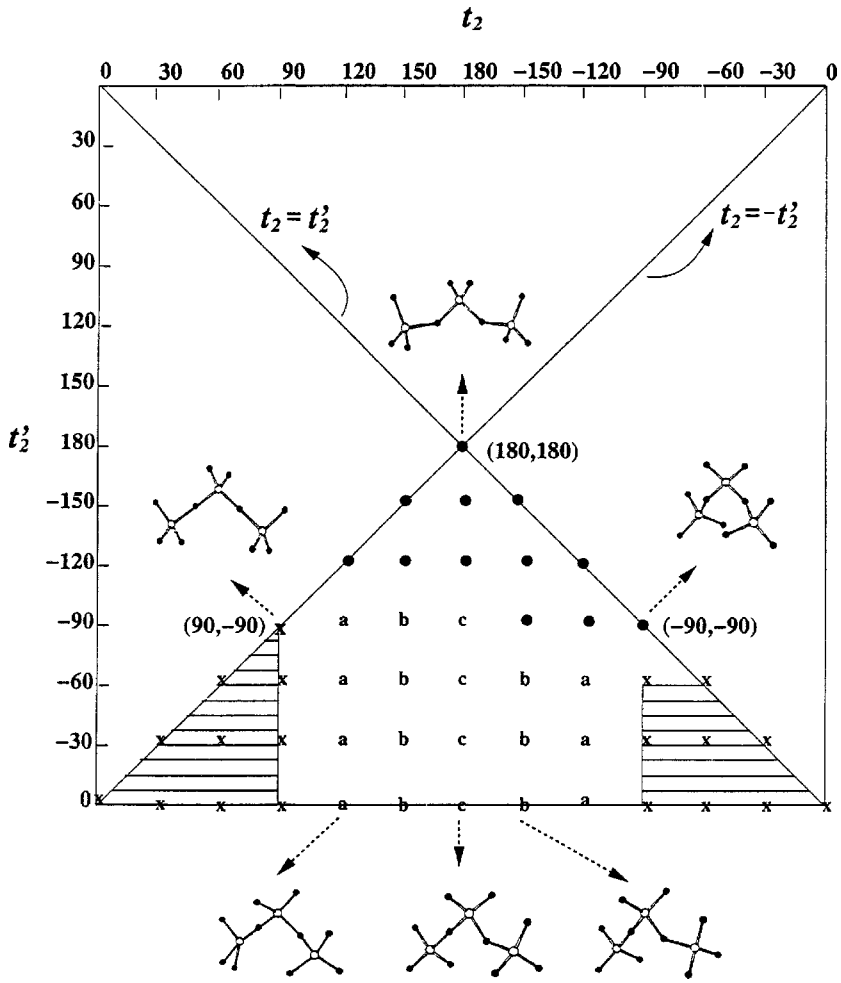


FIGURE 4. The distinct structures (grid points) on the (t_2, t_2') conformational space of TP. "*a*," "*b*," and "*c*" represent, respectively, three differing structures for which one of the two P—O—P angles is linearized (i.e., 180°), whereas the shaded area denotes indistinguishable structures because their P—O—P angles are both 180°. Representative structures taken from the *stg–stg* group for key conformations on the grid surface are shown.

complete linearization on both of the two P—O—P angles, whereas "a," "b," and "c" linearize only one of the two. Based on the results obtained for the sample structures calculated in Table III, it appears that the same pattern of structural degeneracy given by Figure 4 applies to all of the six conformational groups, in that when t_2 or t_2' is smaller than 90° (or equal to 90° when t_2 and t_2' have opposite signs) the structure would degenerate. Consequently, of the 49 (t_2, t_2') grid points only a reduced set can be uniquely defined, yielding, thereby, 16, 35, 60, 25, 35, and 16, as the number of distinct rotational isomers for group *stg–stg*, *stg–skw*, *skw–skw*, *stg–eclp*, *skw–eclp*, and *eclp–eclp*, respectively. Together, the net number of distinct TP rotational isomers is reduced to 187, from the 20,736 (12^4) possible structures that can be generated for a molecule of four rotatable bonds by a 30° torsional search. Although attempts were not made to optimize them all, a selected number, 59, sampling a wide range of conformational energies (ΔE up to 15 kcal/mol, see Table III), appear to be sufficient for characterizing the TP conformational space.

Discussion

RELATIVE STABILITIES OF PHOSPHATE CONFORMATIONS

DP

Table I shows that there is an increase of ~0.5 kcal/mol in energy (ΔE) from *stg* to *skw*, and again from *skw* to *eclp* structures. The ordering of

TABLE III.
Conformational Energy Differences (ΔE, in kcal/mol) and P—O*—P Angles (deg) of Distinct Triphosphate (TP) Rotational Isomers[a] (HF/6-31G* Results).

(t_2, t_2')[b]	P_1—O*—P_2	P_2—O*—P_3	ΔE[c]
stg–stg			
(180, 180)	146.0	146.0	0.00
(180, −150)	145.5	147.7	0.86
(−150, −150)	146.2	146.2	1.00
(150, −150)	148.7	148.4	2.62
(−**150**, −**120**)	144.2	150.3	**2.74**
(180, −120)	144.1	153.9	3.02
(180, −90)	142.7	180.0	4.97
(**150**, −**120**)	147.7	158.0	**5.35**
(150, −90)	144.3	180.0	6.22
(−120, −90)	146.4	154.6	7.50
(120, −90)	150.5	180.0	9.51
(−**90**, −**90**)	154.2	154.2	**11.2**
(90, −90)	180.0	180.0	12.8
stg–skw			
(**180**, **180**)	146.0	146.5	**0.64**
(180, −150) I	145.9	147.6	1.41
(180, −150) II	145.1	150.6	1.90
(150, −150)	151.3	148.1	3.68
(180, −120)	143.8	156.1	3.78
(180, −90)	142.6	180.0	5.44
(−150, −90)	142.9	160.1	5.80
(150, −90)	144.3	180.0	6.56
(−120, −90)	146.1	155.7	8.07
(120, −120)	160.8	155.5	9.53
(−90, −90)	155.5	153.6	11.8
(90, −90)	180.0	180.0	13.2
skw–skw			
(**180**, **180**) I	146.4	146.4	**1.27**
(180, 180) II	146.4	146.6	1.41
(−150, −150)	146.7	146.7	2.35
(−150, −120) I	146.4	151.9	4.45
(−150, −120) II	144.7	152.5	4.54
(180, −120)	144.2	156.8	4.73
(150, −120)	149.2	159.7	7.30
(150, −90)	146.6	180.0	7.77
(120, −90)	152.5	180.0	10.8
(90, −90)	180.0	180.0	14.1
stg–eclp			
(**180**, **180**)	145.9	147.3	**1.34**
(180, −150) I	146.7	147.5	2.18
(180, −150) II	145.1	151.2	2.46
(−150, −150)	149.5	145.8	2.70
(−150, −120)	147.6	150.0	4.61
(180, −90)	142.6	180.0	5.89
(−150, −90) I	143.1	161.8	6.38
(−150, −90) II	146.3	158.9	7.19
(150, −90)	147.6	180.0	8.10
(−90, −90)	155.0	157.8	12.6
(90, −90)	180.0	180.0	13.8
skw–eclp			
(**180**, **180**)	146.3	147.2	**2.03**
(−150, −150)	149.6	146.7	3.57
(150, −150)	150.5	151.3	5.26
(90, −90)	180.0	180.0	14.6
eclp–eclp			
(**180**, **180**)	147.0	147.0	**2.78**
(180, −150)	146.2	150.8	3.93
(−150, −150)	149.1	149.1	4.58
(150, −120)	149.4	160.1	8.27
(−150, −90)	146.0	161.8	8.59
(150, −90)	147.1	180.0	9.27
(120, −120)	159.1	159.1	11.4
(−90, −90)	158.3	158.2	14.5
(90, −90)	180.0	180.0	15.4

[a] See Figure 1 for the labeling of atoms and torsions.
[b] I and II represents two different structures with identical (t_2, t_2') values.
[c] With respect to the extended, (180,180) conformation of the *stg–stg* group, which is an energy minimum.

these ΔEs of DP structures is consistent with electrostatic repulsion acting as a major factor in the conformational stabilities of electron-rich pyrophosphate anions. The effect of electrostatic repulsion is manifested in the opening of the P—O*—P angle from 159° in *stg* structures to about 165° in *eclp* structures (Table I). These ΔEs also imply that the entire torsional space of DP is readily accessible, because not only structures within the same group are separated by a negligible amount of energy ($\Delta E <$ 0.1 kcal/mol), but also structures between *eclp* and the other two groups differ by less than 1.2 kcal/mol (Table I). Consequently, the (t_1, t_1') potential surface of DP is very flat (Fig. S3), such that within the staggered arrangement viewed along the P—P axis, the bridge oxygen (O*) could swing a full circle without being impeded by any significant barrier. Indeed, the lowest vibrational mode, of 14 cm^{-1} harmonic frequency, for DP *stg* I corresponds to a swinging motion of the bridge oxygen (Fig. S8), which is a movement confined within any of the several troughs of the potential energy surface (Fig. S3).

The *ab initio* calculations thus suggest that although essentially the entire potential surface of DP is accessible, intrinsically the stagger form would be somewhat favored. This is in agreement with an observation made long ago on conformations of biological and inorganic pyrophosphate ions.[22, 23] The

all-*trans* eclipsed form (*eclp* I), which is known to be especially suited for bidentate complexation with metal ions,[12] is calculated to be only 1.09 kcal/mol above *stg* I, the global minimum (see Table I and Fig. S1).

MDP

It can be seen from Table S2 that all of the computed structures of the MDP phosphate conformations are separated by less than 2.5 kcal/mol, with many of them by less than 0.5 kcal/mol. This indicates that the potentially interfering methyl group is not very inhibitory, because a wide range of phosphate conformations can still be accessed. However, compared to DP, *gauche* structures (t_2 within $60 \pm 30°$) are evidently favored in each of the three MDP phosphate rotational groups. This is better seen in Figure 3(A) where the (t_1, t_2) energy surface of MDP is shown. Namely, at any t_2, the potential profile of t_1 torsion is similar to that in DP (Fig. S3), exhibiting a pattern that would be expected for a symmetrized (or pseudosymmetrized) threefold rotation. By comparison, that of t_2 is altered somewhat by the favoring of *gauche* structures, which results in a mild barrier, ~2.5 kcal/mol, at the *trans* conformation.

TP

With one more excess electron than DP and two more than MDP, TP clearly favors extended conformations where the electrostatic repulsion can be better dissipated. Thus, in all of the six groups, the extended, *anti–anti* conformation of (t_2, t_2') being $(180°, 180°)$ is most favored, and in general, for the same (t_2, t_2') torsions those belonging to the more crowded group have higher ΔEs (e.g., for the $(-150°, -150°)$ isomers, *stg–stg* is more stable than *skew–skew* than *skw–eclp* than *eclp–eclp*; see Table III). This conformational preference for avoiding electrostatic repulsion is manifested both in the linearization of more folded structures, as noted above, and in the energetic differences between different structures. For example, in the *stg–stg* group, the structure of $(150°, -120°)$, having t_2 and t_2' in opposite signs, and consequently its two terminal phosphates folding toward each other, is 2.61 kcal/mol higher in energy than the structure of $(-150°, -120°)$ (ΔE 5.35 vs. 2.74, see Table III). Moreover, it appears that the relative energies are roughly additive for the extended structures. Thus, the $(180°, 180°)$ structures of the *stg–skw* and the *stg–eclp* group are 0.64 and 1.34 kcal/mol higher in energy than their *stg–stg* counterpart, respectively. The increments in ΔE of these three *stg-X* (X = *stg*, *skw*, *eclp*) TP structures going from *stg* to *skw*, 0.64, and then to *eclp*, 0.7, are only slightly larger in magnitude than the ~0.5 kcal/mol mentioned above in the corresponding DP structures. Similarly, at $(180°, 180°)$, ΔE of *skw–skw*, 1.27, is about twice that of *stg–skw* ($0.64 + 0.64 = 1.28$); ΔE of *skw–eclp*, 2.03, is close to the 1.98 that the ΔE of *stg–skw* and of *stg–eclp* added together equal; and ΔE 2.78 of *eclp–eclp* is also close to twice that (1.34) of *stg–eclp*. Although these energy analyses have stretched the interpretation on the accuracy of the *ab initio* method (~1 kcal/mol), taken as a whole and in view of the more meaningful energetic trends in a qualitative sense, these results support the notion that rotations of the two terminal phosphate groups do not interfere with each other in the extended conformations, and that they are quite flexible; in fact, very much similar to what was found in DP. On the other hand, and in contrast to what was found in MDP, rotations of the central two torsions, t_2 and t_2', are highly confined such that a folded conformation like the $(-90°, -90°)$ structure of the *stg–stg* group is as high as 11.2 kcal/mol in energy above the extended $(180°, 180°)$ structure of the same group (see Fig. 4 for the depiction of the two structures). As may be seen from Table III, similar magnitudes in this energy difference persist in the other five groups.

The extended TP structure of the *stg–stg* group was further shown to be an energy minimum in a full-degree optimization followed by calculation of vibrational frequencies. As discussed above, the central two torsions are energetically restricted to the *anti–anti* conformation, and in this conformation the terminal phosphates would behave independently to favor, albeit marginally, a staggered arrangement, like DP (and, to a lesser extent, the phosphate conformation of MDP). On these grounds, it is likely that this structure, a conformation of C_{2v} symmetry (Fig. S5), is the global minimum for TP in isolation. Consistent with this conjecture, the crystal structure of sodium triphosphate ($Na_5P_3O_{10}$) was found to adopt this conformation,[24] although its torsions (165° for both) deviate somewhat from a perfect *trans* angle due, presumably, to ionic and/or crystal packing forces. Likewise, previous calculations[3, 15] on TP were done on this conformation.

MDP METHYL CONFORMATION: sp3 C—H...O HYDROGEN BOND AND *GAUCHE* EFFECT

It can be seen from Table S3 that, like the phosphate rotational isomers (Table S2), *gauche* confor-

mation is generally favored in the methyl rotational isomers regardless of whether the structure is a *stg*, *skw*, or *eclp* type. But unlike the phosphate case, in all of the three methyl conformational groups non-*gauche* structures are notably higher in energy, especially for those with t_3 equal to or larger than 120°. These non-*gauche* methyl conformations are appreciably less stable, owing to a combination of two factors discussed as follows.

The first factor is a diminishing intramolecular hydrogen bonding between the methyl group and the terminal, unmethylated phosphate group. With a separation as short as 2.1 and 2.2 Å in many low-energy MDP structures (see both Table S2 and Table S3), this intramolecular sp3-hybridized C—H. . .O hydrogen bond, while unconventional, cannot be discounted as being of only marginal strength. In fact, its influence on the stability of MDP conformers is quite evident, in that those of a lower ΔE tend to have a shorter C—H. . .O bond, which is typified in the surprising stability of the *cg* structure discussed above (Table II and Fig. S4). Indeed, Mulliken population analysis showed that in the structures where this unusual hydrogen bond is facilitated, the participating hydrogen atom is significantly more polarized (see the derived partial charges shown in Fig. S4). Note that while the methyl may also form an intramolecular hydrogen bond with the adjacent phosphate group, in which case H and O will be separated by four covalent bonds instead of six, a stronger steric constraint thereby embodies[25] would likely render such arrangements unfavorable.

The existence of C—H. . .O hydrogen bonding is gaining attention in recent years,[26, 27] with a growing body of crystallographic evidence supporting a nonnegligible role in structure and function of small organics,[27] carbohydrates,[25] proteins,[28–30] and nucleic acids.[12, 31–35] Our *ab initio* results further suggest that the C5′ methylene group of nucleotides, which takes the same position as MDP's methyl, can be significantly polarized to form, for example, intermolecular hydrogen bonds with electronegative atoms such as carbonyl oxygen of binding proteins, thereby contributing to protein–nucleotide recognition. Indeed, evidence from surveying a set of diverse and nonredundant NAD– and NADP–protein complex structures determined by X-ray crystallography indicates that the proposed C5′—H. . .O hydrogen bond is widely present and conserved in various modes of protein–dinucleotide binding.[36]

The second factor that contributes to non-*gauche* MDP methyl conformations being of much higher energy is the disruption of a stereospecific interaction favoring *gauche* torsions, i.e., the "*gauche* effect."[37] A strong *gauche* effect in a methyl-substituted phosphate compound, 1,2-dimethyl phosphate, has been reported and clearly analyzed.[20] In MDP, the *gauche* effect is compounded by the formation of a C—H. . .O intramolecular hydrogen bond. As a result, the *gauche* effect due to the methyl group (t_3 torsion) is more profound than that due to the phosphate group (t_2 torsion), as non-*gauche* conformations are of considerably higher energy in the former than in the latter (e.g., the *stg* structure at $t_3 = 180°$ in Table S3 is more than 5 kcal/mol higher in energy than its counter structure in Table S2). This is further illustrated in Figure 3(B), where the landscape of the energy surface is drastically pronounced by an elevated barrier height (as high as ~8 kcal/mol) separating *gauche* conformers. In other molecules that exhibit gauche effect, the orbital interaction origin of the *gauche* effect has been explored with the natural bond orbital (NBO) analysis,[37b, 38] and the electrostatic origin of the gauche effect has been analyzed on the basis of dipole–dipole, dipole–quadrupole, and quadrupole–quadrupole interactions.[39] Similar analyses may be applied to elucidate the origin of the gauche effect exhibited by MDP.

GEOMETRIES OF THE PYROPHOSPHATE LINKAGE AND FLEXIBILITY OF P-O-P ANGLES

The calculated values for the P—O* bond and the P—O*—P angle of the minimum-energy structure of DP, *stg* I, are 1.667 Å and 159°, respectively. These values are much larger than the corresponding statistical average, 1.60 Å and 133°, of experimentally determined crystal structures of various inorganic pyrophosphates.[6, 12] These differences may reflect a general property about the nature of pyrophosphate (and, to a lesser extent, phosphate) compounds; namely, their electronic structures can be greatly influenced by environmental forces that were not considered in the present calculations. These environmental forces may include substitution (e.g., protonation and methylation), solvation, metal complexation, crystal packing, etc. Thus, for example, in MDP the addition of the methyl group causes the farther pyrophosphate bond, P—O*, to stretch, and the near bond, P*—O**, to compress (Table S2, ~1.76 and ~1.56 Å, respectively, vs. 1.67 Å in DP). Similar bond variations can also be seen in the reported optimized structures of protonated anions of pyrophosphoric acid.[4, 9] Crystal field[40] and the resulting multiple hydrogen bonding[41] have been evoked to explain the shortening of peptide

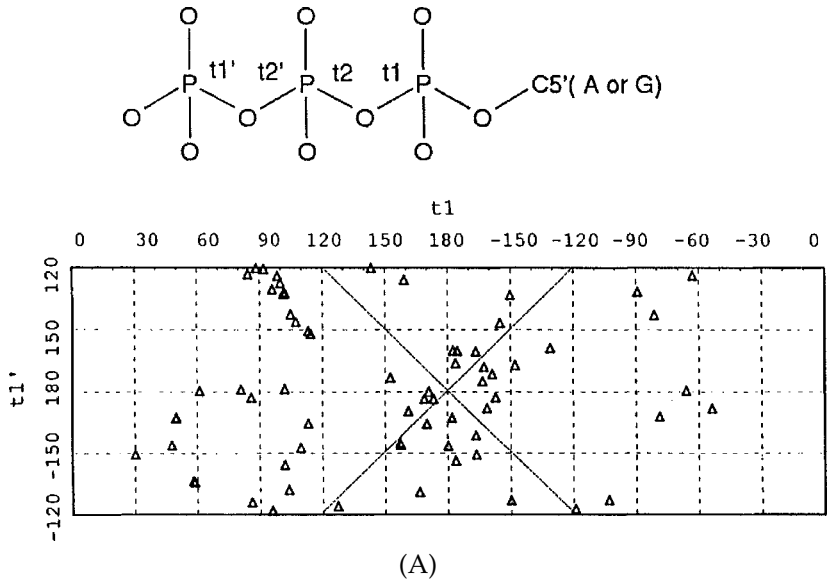
bond from gas phase (~1.37 Å) to crystal phase (~1.32 Å). Similar forces may exist to significantly alter pyrophosphate bond lengths from an isolated to a condensed state. The most dominating force to modify the geometry of pyrophosphates in different states is probably that of solvation and metal complexation, however. In a recent study combining analysis of 178 crystal structures of phosphate compounds (not including pyrophosphates) and *ab initio* calculations of model systems, Schneider et al.[21] showed that metal complexation is an influential modulator of the phosphate geometry, and that in many −2 charged phosphates their interaction with metal captions is mediated by water molecules. For pyrophosphates, solvation, metal complexation, and protein binding are likely to exert an even greater geometry-modifying effect.

The P—O—P angle, excluding those linearized, is about 150° for MDP structures belonging to both the *stg* and the *skw* group, and 157° on average for those to the *eclp* group (see Table S2). Therefore, the two phosphate groups are significantly more bent, by about 10°, in MDP than in DP, a geometric consequence that may be attributed to one less charge, thus less severe electrostatic repulsion, in MDP, as well as to an electrostatic attraction due to the C—H. . .O hydrogen bonding.

With a blunt angle much wider than the tetrahedral standard of 109°, the P—O—P angle is considered to be extremely floppy, which may, in fact, play a role in the unusual flexibility of the pyrophosphate linkage. Indeed, normal mode analysis using harmonic approximation showed that the vibrational frequency of P—O—P bending is quite low, among the lowest, for each of the three model compounds (Fig. S6; a complete listing for all of the vibrational modes is supplied in Fig. S8). Consistent with the vibrational analysis, further calculations showed that with HF/6-31G* only 0.3 kcal/mol is required to completely straighten out this angle for DP *stg* I (with HF/6-311++G* no energy is required because global minimum calculated with the larger basis set has a linearized P—O—P angle[9]). Flexibility may be an intrinsic property of X—O*—X type angles (where X is a relatively large element such as phosphorus), as an extremely flexible Si—O*—Si (or Al) angle in silicates and aluminosilicates has also been noted.[42] As such, transitions between rotational isomers of pyrophosphate molecules such as DP can be achieved via twisting or inverting the P—O*—P angle instead, or most likely through a coupling between this angle and the dihedral of the P—O* bond rotation.

COMPARISON WITH OBSERVED PYROPHOSPHATE CONFORMATIONS

Essentially, all ranges of phosphate conformations have been observed in crystal structures of both organic and inorganic pyrophosphate species.[11–13] Therefore, our *ab initio* calculations are in accord with the experimental observations in reflecting a flexible pyrophosphate bond rotation. However, the rigidity of TP's central two torsions as revealed by the *ab initio* calculations is at variance with the free rotation notion constantly evoked to describe the conformational flexibility of the pyrophosphate linkage. To some degree the disparity may have been a result of nonspecificity in previous surveys[11–13] wherein terminal and nonterminal phosphates were not distinguished. To test this conjecture, we examine 42 protein-ATP (or GTP) complex structures from the Protein Data Bank (PDB)[43] for the distribution of the lingand's four pyrophosphate torsions. The results, shown in Figure 5(A) and 5(B), revealed that in these protein-bound ATPs and GTPs (a total of 70 of them) a full range of torsional angles is sampled for the terminal torsion t'_1, while rotation of the phosphate attaching adenosine or guanosine (i.e., t_1 torsion) is only slightly restricted—prohibited at highly eclipsed conformations [between −30 and 30° torsions, Fig. 5(A)]. In comparison, the central two torsions, t_2 and t'_2, and t_2 in particular, are less distributed, and they are further refrained from the eclipsed conformations [Fig. 5(B)]. It is notable, however, that many (t_2, t'_2) torsions in these protein-bound adenine and guanine triphosphates have shifted from the extended (180°, 180°) conformation to folded conformations whose ΔE would amount to greater than 10 kcal/mol on the *ab initio* energy surface of an isolated triphosphate anion as indicated by Figure 4 in conjunction with the energy data of Table III. In fact, unlike the flanking torsions [Fig. 5(A)], the two central torsions of the ATPs and GTPs surveyed do not cluster at all at the fully extended, (180°, 180°), conformation [Fig. 5(B)]. In line with the finding of Schneider et al.'s study on phosphates,[21] the apparent disparity between Figure 4 and Figure 5(B) may, therefore, suggest that the conformational preferences of the pyrophosphate linkage are highly regulable in protein-bound complexes wherein solvation and metal bindings are a norm. Indeed, results of further *ab initio* calculations, presented in Table S4, indicated that the stability for certain experimentally observed pyrophosphate torsions of TP can be significantly enhanced by solvation.



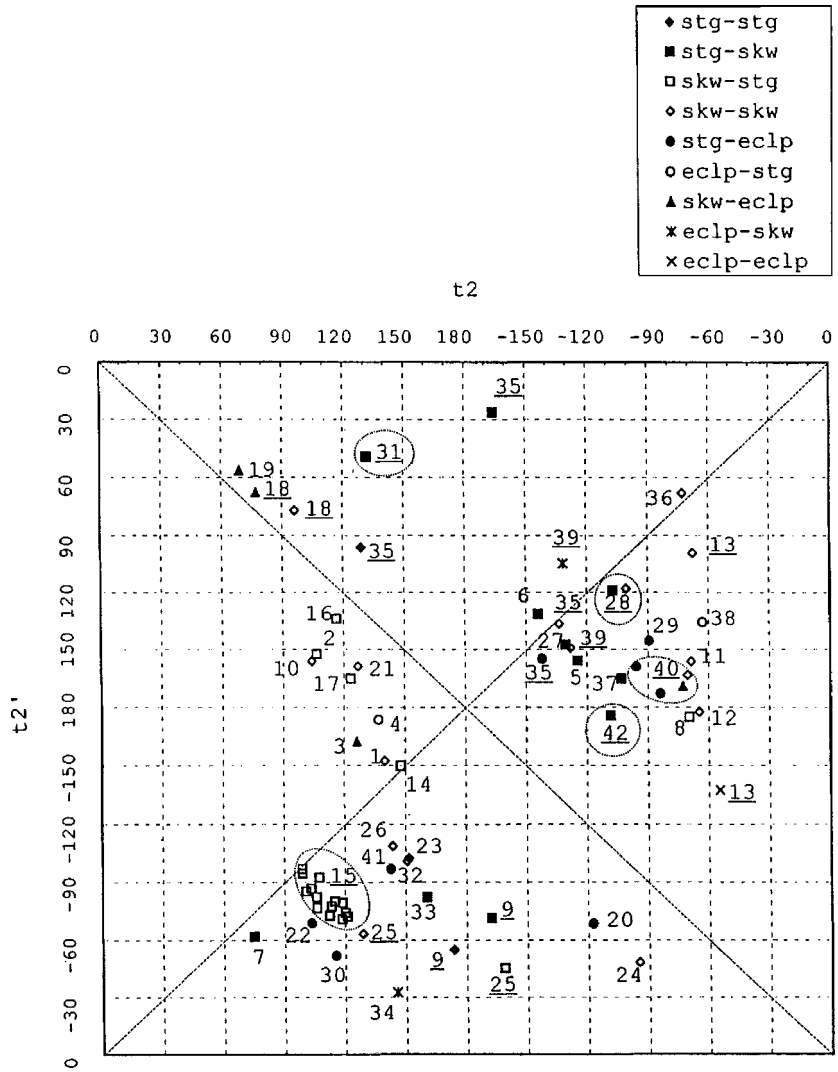
(A)

FIGURE 5. The pyrophosphate torsional angles as observed in crystal structures of protein-bound ATP and GTP: (A) flanking torsions (t_1 and t_1', t_1 being the torsion of the phosphate attaching to the C5′ atom of adenosine or guanosine; note that because of the threefold rotational symmetry, only a third, the one closest to a *trans* angle, of t_1' torsions are shown), (B) central torsions (t_2 and t_2'). These torsions are presented in the same way as in Figure 4, with the diagonal lines indicating symmetrical relationships. PDB codes of the 42 complex structures surveyed, as numbered in the order of decreasing resolution, are: 1:1kax, 2:1kay, 3:1kaz, 4: 2bup, 5:1ayl, 6:1aq2, 7:1hck, 8:1csn, 9:1nsy(2), 10:1a60, 11:1atp, 12:1phk, 13:1fin(2), 14:1ngf,15:1der(14), 16:1nge, 17:ngg, 18:1ckm(2, GTP), 19:1ckn(GTP), 20:1gtr, 21:ngh, 22:3pgk, 23:2btf, 24:1a0i, 25:1jst(2), 26:1qrs, 27:2rap(GTP), 28:4at1(2), 29:521p(GTP), 30:1hlu, 31:1bcp(2), 32:1qrt, 33:1atn, 34:1gol, 35:1kmn(4), 36:1mab, 37:1plk(GTP), 38:7at1, 39:1asz(2), 40:1nbm(4), 41:1qru, 42:3r1r(3). These structures were selected by a resolution cutoff of 3 Å, the first nine equal to or better than 2 Å. Except noted in the parenthesis following the PDB code, there is a single ATP occurrence in the complex structure. Those with multiple occurrences of ATP or GTP can also be identified in (B), as they are underlined. As may be seen, whereas some multiple occurrences produce very similar pyrophosphate torsions (complex 15, 28, 31, 40, and 42; circled), others do not (e.g., complex 13). Also, all of the conformational groups defined in Figure S2 and Table III are sampled (see symbol legends).

The influence of *gauche* or anomeric effect on the orientation of the P—O bond in both nucleotides[12, 13] and lipids[20, 44] has been well documented. That the *gt* structure of MDP is calculated here to be as much as 6.55 kcal/mol higher in energy than the lowest energy *gg* I structure (see Table II) is a reconfirmation of this effect.

Conclusion

Using *ab initio* methods of mostly HF/6-31G* with a few cases additionally treating the effects of correlation, diffusion, and solvation, we have characterized the gas-phase potential energy surface pertaining to the bond rotation of the pyrophosphate P—O—P linkage as embodied in three model species, diphosphate ($P_2O_7^{-4}$), methyl diphosphate ($CH_3P_2O_7^{-3}$), and triphosphate ($P_3O_{10}^{-5}$). Our main findings are the following: 1) the generally perceived free rotation of the phosphate group in pyrophosphates is essentially borne out by the *ab initio* calculations, which suggest further that the extreme flexibility of the pyrophosphate linkage may be greatly assisted by a very floppy P—O—P angle. 2) The addition of a methyl group can elicit a drastic change on the electronic structure of pyrophosphate—for the conformations of methyl diphosphate studied here, the methyl addition results in a profound *gauche* effect and a rather significant sp3-hybridized C—H. . .O hydrogen bond. 3) In contrast to terminal phosphates, the central phosphate of TP requires considerable energy to rotate. This restriction in pyrophosphate rotation is ivery profound in the *ab initio*-calculated energy surface of the triphosphate model, but is subdued greatly in protein-bound ATP and GTP. It is suggested that environmental forces such as solvation and metal complexation contributed to lower the rotational barriers of the pyrophosphate linkage. Altogether,



(B)

FIGURE 5. (Continued)

our *ab initio* calculations addressed several aspects of the molecular forces governing the conformational properties of the pyrophosphate linkage and, for the first time, detailed the potential surfaces of the three models. These data should be valuable to the derivation and validation of empirical force fields for simulating the pyrophosphate functionality.

Acknowledgments

We are grateful to Dr. H. H. Chang for assisting in much of the *ab initio* work, and for very critically reading an earlier version of the manuscript. We would like to thank Mr. H. Y. Chang for his involvement in the initiation of this project. The computer center of Academia Sinica and the National Center for High-performance Computing (NCHC) are acknowledged for providing the needed computational resources.

Supplementary Materials

There are four tables (Table S1–S4) and eight figures (Figure S1–S8).

References

1. Martin, D. W.; Mayes, P. A.; Rodwell, V. W. Harper's Review of Biochemistry; Lange Medical Publications: Los Altos, CA, 1983.

2. (a) Oesper, P. Arch Biochem Biophys 1950, 27, 255; (b) Hill, T. L.; Morales, M. F. J Am Chem Soc 1951, 73, 1656; (c) Fukui, K.; Morokuma, K.; Nagata, C. Bull Chem Soc Jpn 1960, 33, 1214; Pullman, B.; Pullman, A. Quantum Biochemistry; Interscience: New York, 1963.
3. Boyd, D. B.; Lipscomb, W. N. J. Theor Biol 1969, 25, 403.
4. Ma, B.; Meredith, C.; Schaefer, H. F., III J Phys Chem 1994, 98, 8216.
5. Hayes, D. M.; Kenyon, G. L.; Kollman, P. A. J Am Chem Soc 1978, 100, 4331.
6. O'keeffe, M.; Domenges, B.; Gibbs, G. V. J Phys Chem 1985, 89, 2304.
7. Ewig, C. S.; Van Wazer, J. R. J Am Chem Soc 1988, 110, 79.
8. Saint-Martin, H.; Ortega-Blake, I.; Les, A.; Adamowicz, L. Biochim Biophys Acta 1991, 1080, 205.
9. Colvin, M. E.; Evleth, E.; Akacem, Y. J Am Chem Soc 1995, 117, 4357.
10. Lipmann, F. Adv Enzymol 1941, 1, 99.
11. Saenger, W.; Reddy, B. S.; Mühlegger, K.; Weimann, G. Nature 1977, 267, 225.
12. Saenger, W. Principles of Nucleic Acid Structure; Springer Verlag: New York, 1984.
13. Moodie, S. L.; Thorton, J. M. Nucleic Acids Res 1993, 21, 1369.
14. Cannon, J. F. J. Comput Chem 1993, 14, 995.
15. Laki, K.; Seel, M.; Ladik, J. J. Theor Biol 1977, 67, 489.
16. Pavelites, J. J.; Gao, J.; Bash, P. A.; Mackerell, A. D., Jr. J Comput Chem 1997, 18, 221. This article published a newly derived CHARMM force field for nicotinamide adenine dinucleotides and the pyrophosphate linkage of nucleotides, in which a number of MDP rotational isomers were optimized by varying the t_2 torsion (P—O*—P*—O**) with HF/6-31+G*. The authors reported a flat t_2 surface at gauche methyl conformations (t_3), which is consistent with our findings, but did not mention the C—H. . .O hydrogen bond.
17. Gaussian 94, Revision B.2, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Pittsburgh, PA, 1995.
18. (a) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A., III; Honig, B. J Am Chem Soc 1994, 116, 11875; (b) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D; Honig, B. J Phys Chem 1996, 100, 11775.
19. Jaguar v3.5, Schrödinger, Inc.: Portland, OR, 1998.
20. Liang, C.; Ewig, C. S.; Stouch, T. R.; Hagler, A. T. J Chem Soc 1993, 115, 1537.
21. Schneider, B.; Kabelac, M.; Hobza, P. J Am Chem Soc 1996, 118, 12207.
22. Pletcher, J.; Sax, M. Science 1966, 154, 1331.
23. Sundaralingam, M. Biopolymers 1969, 7, 821.
24. Cruickshank, D. W. J. Acta Crystallogr 1964, 17, 674.
25. Steiner, T.; Saenger, W. J Am Chem Soc 1992, 114, 10146.
26. Wahl, M. C.; Sundaralingam, M. Trends Biochem 1997, 22, 97.
27. Desiraju, G. R. Acc Chem Res 1996, 29, 441.
28. (a) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. J Mol Biol 1994, 241, 83; (b) Derewenda, Z. S.; Lee, L.; Derewenda, U. J Mol Biol 1995, 252, 248.
29. Bella, J.; Berman, H. M. J Mol Biol 1996, 264, 734.
30. Chakrabarti, P.; Chakrabarti, S. J Mol Biol 1998, 284, 867.
31. Jeffrey, G. A.; Saenger, W. Hydrogen Bonding in Biological Structures; Springer Verlag: Berlin, 1991.
32. Rao, S. T.; Sundaralingam, M. J Am Chem Soc 1970, 92, 4963.
33. Leonard, G. A.; McAuley-Hecht, K.; Brown, T.; Hunter, W. N. Acta Crystallogr 1995, D51, 136.
34. Mayer-Jung, C.; Moras, D.; Timsit, Y. J Mol Biol 1997, 270, 328.
35. Mandel-Gutfreund, Y.; Margalit, H.; Jernigan, R. L.; Zhurkin, V. B. J Mol Biol 1998, 277, 1129.
36. Chu, P.-Y.; Hwang, M.-J. J Mol Biol 1998, 279, 695.
37. For examples, see (a) Kirby, A. J. The Anomeric Effect and Related Stereoelectronic Effects at Oxygen; Springer Verlag: Berlin, 1983; (b) Reed, A. E.; Schleyer, P. v. R. J Am Chem Soc 1987, 109, 7362.
38. Salzner, U.; Schleyer, P. v. R. J Am Chem Soc 1993, 115, 10231.
39. Chang, Y.-P.; Su, T.-M. J Mol Struct (Theochem) 1996, 365, 183; and references therein.
40. Popelier, P.; Lenstra, A. T. H.; Van Alsenoy, C.; Geise, H. J. J Am Chem Soc 1989, 111, 5658.
41. Schafer, L.; Newton, S. Q.; Cao, M.; Peeters, A.; Van Alsenoy, C.; Wolinski, K.; Momany, F. A. J Am Chem Soc 1993, 115, 272.
42. Hill, J.-R.; Sauer, J. J Phys Chem 1995, 99, 9536.
43. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. E., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. J Mol Biol 1977, 112, 535.
44. Hauser, H.; Pascher, I.; Sundell, S. In R. Brasseur, Ed., Molecular Description of Biological Membranes by Computer Aided Conformational Analysis; CRC Press: Boston, MA, 1990.